An Initial Intercept Iteratively Adjusted (IIIA) Controller: An Enhanced Double EWMA Feedback Control Scheme

Sheng-Tsaing Tseng, *Member, IEEE*, Wheyming Song, and Yu-Chi Chang

***Abstract*—The double exponentially weighted moving average (dEWMA) controller is a popular run-to-run controller for a drifted process. It uses the information collected from past process data to adjust the process parameters for the later runs. The dEWMA controller can guarantee long-term stability under suitably fixed discount factors and fairly regular conditions. However, it usually requires a large number of runs to bring the process output to meet its target and, thus, may leave the output out of its specification at the beginning of the first few runs. Hence, dEWMA controller is not suitable for the processes with short production runs. To reduce the possibility of high rework rate, we propose an enhanced dEWMA controller, which we refer to as the initial intercept iteratively adjusted (IIIA) controller, to further eliminate the off-target (nonrandom biases) of the process output. We derive an analytic expression of the process output of the IIIA controller, and present several examples to show that the IIIA reduces the process mean square error significantly, as well as the rework rate.**

***Index Terms*—Double exponentially weighted moving average (dEWMA) controller, initial intercept iteratively adjusted (IIIA) controller, recursive least squares (RLS), run to run (R2R) process control.**

I. Introduction

STATISTICAL process control (SPC) and engineering process control (EPC) are widely used to monitor and adjust industrial processes ranging from part manufacturing to process manufacturing. A general introduction to SPC and EPC techniques can be found in [2], [7], [15], and [16]. During the last decade, various integrated approaches have been proposed to combine the advantages of SPC and EPC [14]. In particular, for semiconductor manufacturing, [9] proposed an efficient model-based process control method, which they named run-to-run (R2R) feedback control [5].

The R2R feedback control focuses on how to adjust the process recipe so that the process output can be adjusted to a desired target for every production run. One typical example is photolithography of a semiconductor manufacturing process, for which the R2R approach is usually applied to control the feature size (critical dimension) by adjusting parameters such as the exposure dose (energy) on the wafers for each run. Other examples include reactive ion etching (RIE) [8], and chemical mechanical polishing (CMP) of the semiconductor manufacturing processes [4]. In most R2R applications, the exponentially weighted moving average (EWMA) controller (control scheme) has received a great deal of attention. Reference [9] proposed a single EWMA controller and investigated the stability and sensitivity of its process output. Reference [3] discussed the application of the double EWMA (dEWMA) controller on polysilicon gate etching process. References [6] and [12] further investigated the stability conditions, the long-run behavior, the transient performance, and the determination of the optimal discount factors of the dEWMA controller.

Although the dEWMA controller with suitable discount factors can guarantee long-term stability under fairly regular conditions, it usually requires a moderately large number of runs to have the process output converge to its target. This may lead to severe consequences, such as a high rework rate. This phenomenon is particularly prevalent in small-batch fabrication in semiconductor manufacturing with the drifted processes. To reduce such a high rework rate, we propose an enhanced dEWMA controller, to be referred to as *the initial intercept iteratively adjusted* (IIIA) controller. The advantage of the proposed IIIA controller is its capability of eliminating the off-target deviation (nonrandom bias) of the process output occurred in the conventional dEWMA controller.

This paper is organized as follows. In Section II, we provide the problem statement and its motivations. In Section III, we propose an IIIA controller. In Section IV, we present the analytical results of the IIIA controller, and demonstrate the performance of the IIIA controller using several examples. In Section V, we extend these results by comparing the performance of the IIIA controller with that of two other controllers (traditional dEWMA and recursive least squares) in a simulation study. Finally, some concluding remarks are given in Section VI.

II. Motivations and Problem Formulation

Given a drifted single-input–single-output (SISO) process, we assume that the input and output variables follow the following linear relationship:

$$Y_t = \alpha + \beta u_{t-1} + \delta t + \eta_t. \tag{1}$$

Here, Y_t is the output of tth production run, α and β are the intercept and slope of the linear model, respectively, u_{t-1} is the input recipe which is determined at the end of $(t-1)$th run, δ is the drift rate, and η_t is the disturbance which can be suitably

Manuscript received August 19, 2003; revised April 20, 2005. This work was supported in part by the National Science Council, Taiwan, R.O.C., under Contract NSC 92-2213-E-007-066.

S.-T. Tseng is with the Institute of Statistics, National Tsing-Hua University, Hsin-Chu 30043, Taiwan, R.O.C. (e-mail: sttseng@stat.nthu.edu.tw).

W. Song and Y.-C. Chang are with the Department of Industrial Engineering, National Tsing-Hua University, Hsin-Chu 30013, Taiwan, R.O.C. (e-mail: wheyming@ie.nthu.edu.tw, yuchi@ctu.edu.tw).

Digital Object Identifier 10.1109/TSM.2005.852114

modeled by a general time series model [1]. If τ denotes the desired performance target value for the process output, then the goal of the feedback control scheme is to adjust u_{t-1} so that the expected output (Y_t) will be as close as τ. The first step of the feedback scheme is to build a prediction model for the process output. Applying design of experiments (DOE), we construct a simple least squares estimate (LSE)

$$\hat{Y} = a_0 + bu \tag{2}$$

where a_0 and b denote the initial estimates of the unknown parameters α and β of (1), respectively. To monitor a linear-drift manufacturing process, [3] proposed the following double exponentially weighted moving average (dEWMA) controller

$$u_t = \left(\frac{\tau - a_t - D_t}{b}\right) \tag{3}$$

where

$$a_t = \lambda_1(Y_t - bu_{t-1}) + (1-\lambda_1)a_{t-1}, \quad 0 < \lambda_1 \le 1 \tag{4}$$

$$D_t = \lambda_2(Y_t - bu_{t-1} - a_{t-1}) + (1-\lambda_2)D_{t-1}, \quad 0 < \lambda_2 \le 1 \tag{5}$$

with $D_0 = 0$. To implement this controller, we need to select suitable discount factors λ_1 and λ_2 to ensure that the following asymptotic stability conditions as satisfied:

$$\lim_{t\to\infty} \mathrm{E}(Y_t) = \tau \tag{6a}$$

and

$$\lim_{t\to\infty} \mathrm{Var}(Y_t) < \infty. \tag{6b}$$

Note that (6a) means that Y_t is asymptotically unbiased, and (6b) implies that the variance of Y_t is asymptotically finite.

References [6] and [12] investigated the above asymptotic stability conditions (6a) and (6b) under a fair general class of process disturbances. It can be shown that, even with rough (inaccurate) estimates of the unknown parameters, the dEWMA controller with suitable discount factors can guarantee the above stability conditions, but it usually requires a moderately large number of runs to bring the process output closer to its target. To overcome this drawback, we propose an enhanced dEWMA controller and show that it can effectively eliminate the off-target bias.

A. Motivation

The goal of this paper is to propose an efficient procedure to ensure that $\mathrm{E}(Y_t) = \tau$ for all $t \ge 3$. We first recall a result from [12].

Result From [12]:

$$\text{For all } t \ge 1, \qquad (Y_t - \tau) = \Gamma_t + W_t \tag{7}$$

where the nonrandom bias Γ_t and the random part W_t can be expressed as follows:

$$\Gamma_t = \Gamma_0 E_{t-1} + \delta F_{t-1} \tag{8}$$

and

$$W_t = \eta_t + \sum_{j=1}^{t-1}(E_j - E_{j-1})\eta_{t-j}. \tag{9}$$

Note that (E_t, F_t) in (8) and (9) can be determined recursively via the following equation:

$$\begin{bmatrix} E_t \\ F_t \end{bmatrix} = \begin{bmatrix} 1-\xi(\lambda_1+\lambda_2) & -\xi\lambda_1\lambda_2 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} E_{t-1} \\ F_{t-1} \end{bmatrix}, \quad \forall\, t = 1, 2, \ldots \tag{10}$$

where $\xi = \beta/b$ and the initial values are $E_0 = 1$ and $F_0 = 0$. Moreover, the initial bias (Γ_0) is given by

$$\Gamma_0 = \alpha + \delta + \xi(\tau - a_0) - \tau. \tag{11}$$

To ensure that $\Gamma_{t+1} = 0$, we require that $a_0 = a_0^*(t)$, where

$$a_0^*(t) = \frac{((\alpha + \delta + (\xi - 1)\tau)E_t + \delta F_t)}{(\xi E_t)}. \tag{12}$$

We call $a_0^*(t)$ the optimal intercept at time t, and the corresponding controller the IIIA controller. Consequently, $\Gamma_{t+1}^* = 0$ and the mean square error (MSE) of the corresponding output Y_{t+1}^* of the $(t+1)$th IIIA process run is

$$\mathrm{E}\left\{\left(Y_{t+1}^* - \tau\right)^2\right\} = \mathrm{Var}(W_{t+1}). \tag{13}$$

The motivation for forcing the off-target bias to be zero, discussed above, is further demonstrated in the following numerical example.

Example 1: Suppose that the parameter settings in (1)–(5) are $\tau = 0$, $\alpha = 4$, $\beta = 2$, $\delta = 0.1$, $b = 2$, $\lambda_1 = 0.1$, $\lambda_2 = 0.45$, and $D_0 = 0$. Suppose further that the disturbance model is ARMA (1,1) with parameters $\phi = 0.9$, $\theta = -0.5$, and $\sigma_\varepsilon = 0.1$. That is

$$\eta_i = 0.9\eta_{i-1} + \varepsilon_i + 0.5\varepsilon_{i-1}.$$

Fig. 1(a) and (b) shows the off-target biases Γ_t (for the dEWMA controller) and Γ_t^* (for the IIIA controller), calculated based on (8), under two different scenarios: initial bias $a_0 = 0$ (i.e., $\Gamma_0 = 4.1$) and $a_0 = 4.1$ (i.e., $\Gamma_0 = 0$). The figure illustrates that, compared with the dEWMA controller; the IIIA controller eliminates the nonrandom bias more efficiently. ■

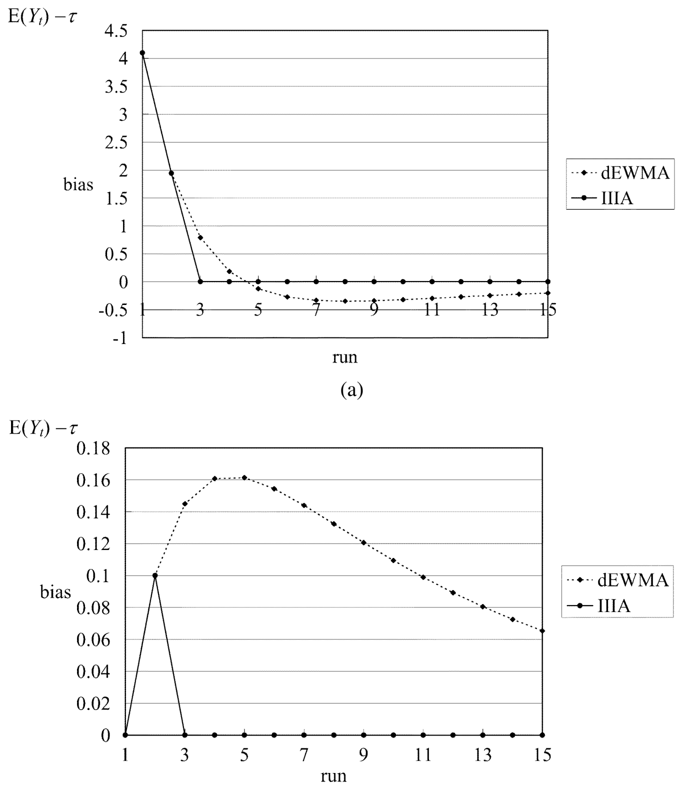
The above results are based on the assumptions that α, δ, and $\xi(= \beta/b)$ in (12) are completely known. In practice, however, these parameters are unknown. Hence, to implement the IIIA controller, we need to consider the following issues more fully.

1) How do we estimate these unknown parameters efficiently?
2) How do we construct the proposed IIIA controller?
3) How do we derive an analytic expression for the IIIA process output?
4) How well does the IIIA controller perform?
5) When will we prefer the IIIA controller to the dEWMA controller?

The above issues will be addressed analytically in Sections III and IV under the assumption that ξ is known. For the case of unknown ξ, we investigate the performance of the IIIA, the dEWMA, and recursive least squares (RLS) controllers via simulation experiments in Section V.

III. Proposed Control Scheme

This section introduces the IIIA controller. Preliminary results, in the form of two lemmas, are given in Part A. Part B contains the procedure of the proposed scheme.



(b)

Fig. 1. (a) Comparison of bias in the IIIA and dEWMA controllers for $\Gamma_0 = 4.1$. (b) Comparison of bias in the IIIA and dEWMA controllers for $\Gamma_0 = 0$.

A. Part A: Two Lemmas

First, we introduce some notation.

1) $$\begin{bmatrix} A_t & O_t \\ B_t & P_t \\ C_t & Q_t \end{bmatrix} = \begin{bmatrix} 1 & \lambda_1 & \lambda_2 \\ 0 & 1-\lambda_1\xi & -\lambda_2\xi \\ 0 & -\lambda_1(\xi-1) & 1-\lambda_2\xi \end{bmatrix} \times \begin{bmatrix} A_{t-1} & O_{t-1} \\ B_{t-1} & P_{t-1} \\ C_{t-1} & Q_{t-1} \end{bmatrix}, \quad t = 1, 2, \ldots$$

2) $$\zeta_t = \sum_{i=0}^{t-1} (\lambda_1 B_i + \lambda_2 C_i)\eta_{t-i}, \quad t = 1, 2, \ldots$$

3) $$\Lambda_t = \sum_{i=0}^{t} (\lambda_1 B_i + \lambda_2 C_i)(t-i), \quad t = 1, 2, \ldots$$

4) $$\upsilon_t = \sum_{i=0}^{t-1} (\lambda_1 P_i + \lambda_2 Q_i)\eta_{t-i}, \quad t = 1, 2, \ldots$$

5) $$\Phi_t = \sum_{i=0}^{t} (\lambda_1 P_i + \lambda_2 Q_i)(t-i), \quad t = 1, 2, \ldots$$

6) $$\begin{bmatrix} K_t \\ L_t \end{bmatrix} = \begin{bmatrix} A_t & B_t & C_t \\ O_t & P_t & Q_t \end{bmatrix} \begin{bmatrix} \tau(\xi-1) \\ a_0 \\ D_0 \end{bmatrix}, \quad t = 1, 2, \ldots$$

where
$$\begin{bmatrix} A_0 & O_0 \\ B_0 & P_0 \\ C_0 & Q_0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Before implementing the IIIA controller, we first need Lemma 1 to estimate α and δ.

Lemma 1: Given (λ_1, λ_2), we have

$$a_t = \alpha A_t + \delta\Lambda_t + K_t + \zeta_t, \quad \forall\, t \geq 1 \tag{14}$$

and

$$D_t = \alpha O_t + \delta\Phi_t + L_t\upsilon_t, \quad \forall\, t \geq 1. \tag{15}$$

The proof, given in Appendix I, can be obtained from (1) and (3)–(5).

The parameters α and δ can be determined by taking the expectations of both sides of the equations in (14)–(15). Note that $\mathrm{E}(\zeta_t) = \mathrm{E}(\upsilon_t) = 0$. Hence

$$\alpha = \frac{[\Phi_t(\mathrm{E}(a_t) - K_t) - \Lambda_t(\mathrm{E}(D_t) - L_t)]}{(A_t\Phi_t - O_t\Lambda_t)} \tag{16}$$

$$\delta = \frac{[O_t(\mathrm{E}(a_t) - K_t) - A_t(\mathrm{E}(D_t) - L_t)]}{(O_t\Lambda_t - A_t\Phi_t)}. \tag{17}$$

Before implementing the IIIA controller, we also need to update the filters a_t and D_t defined in (4) and (5) to reflect the use of $a_0^*(t)$ in place of a_0. If a_t^* and D_t^* denote the corresponding updated filters, then the relationship between (a_t, a_t^*) and (D_t, D_t^*) is given in Lemma 2.

Lemma 2: For fixed (λ_1, λ_2), we have

$$a_t^* = a_t + B_t\left(a_0^*(t) - a_0\right), \quad \forall\, t \geq 1 \tag{18}$$

$$D_t^* = D_t + P_t\left(a_0^*(t) - a_0\right), \quad \forall\, t \geq 1. \tag{19}$$

The proof of Lemma 2 is given in Appendix II.

Now, we are ready to state the proposed IIIA controller.

B. Part B: The IIIA Controller

There are three main phases in the IIIA controller, which is illustrated inside the box with the dashed border in Fig. 2.

Phase I: Estimating α and δ. The estimators of α and δ

$$\hat{\alpha}_t = \frac{[\Phi_t(a_t - K_t) - \Lambda_t(D_t - L_t)]}{(A_t\Phi_t - O_t\Lambda_t)} \tag{20}$$

$$\hat{\delta}_t = \frac{[O_t(a_t - K_t) - A_t(D_t - L_t)]}{(O_t\Lambda_t - A_t\Phi_t)} \tag{21}$$

are obtained by replacing $\mathrm{E}(a_t)$ and $\mathrm{E}(D_t)$ with a_t and D_t in (16) and (17), respectively.

Phase II: Fitting the best a_0. An updated estimator of the intercept $\hat{a}_0^*(t)$ in the predicted model (2) is obtained by replacing α and δ with $\hat{\alpha}_t$ and $\hat{\delta}_t$ in (12), respectively. That is

$$\hat{a}_0^*(t) = \frac{((\hat{\alpha}_t + \hat{\delta}_t + (\xi-1)\tau)E_t + \hat{\delta}_t F_t)}{(\xi E_t)}. \tag{22}$$

Phase III: Updating the filters. Since the new control scheme is implemented sequentially, $a_0^*(t) - a_0$ in (18) and (19) is replaced by $\hat{a}_0^*(t) - \hat{a}_0^*(t-1)$ for all t.

Note that the proposed scheme will be implemented only for $t \geq 3$, since at least two production runs are needed to estimate α and δ. Hence, process output under the proposed IIIA controller is identical to that under the dEWMA controller for $t \leq 2$.

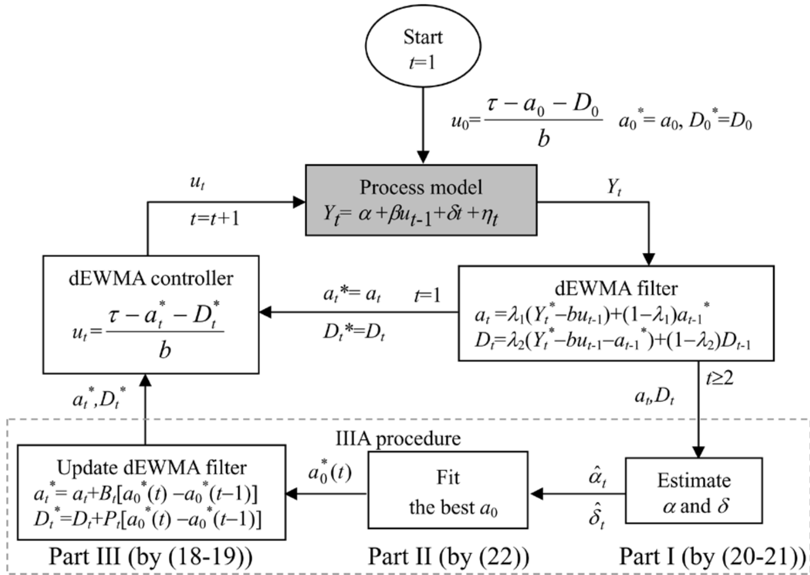


Fig. 2. Diagram of the IIIA controller.

IV. Analytical Results of the IIIA Controller

Three analytical results are presented here. First, $(Y_t^* - \tau)$ is expressed as the sum of two terms, which is then used to compare with $(Y_t - \tau)$ in (7). Let MSE_t and MSE_t^* denote the mean square errors of Y_t and Y_t^*, respectively. Result 2 gives the expression for MSE_t^*. In Result 3, we derive the necessary and sufficient conditions for the inequality $\mathrm{MSE}_t^* < \mathrm{MSE}_t$.

One more notation is defined below. For all t and i, define

$$h_{t-1,i} = \frac{E_{t-1}U_{t-1,i} + (E_{t-1} + F_{t-1})V_{t-1,i}}{(O_{t-1}\Lambda_{t-1} - A_{t-1}\Phi_{t-1})}$$

where

$$U_{t-1,i} = \Phi_{t-1}(\lambda_1 B_i + \lambda_2 C_i) - \Lambda_{t-1}(\lambda_1 P_i + \lambda_2 Q_i)$$

and

$$V_{t-1,i} = -O_{t-1}(\lambda_1 B_i + \lambda_2 C_i) + A_{t-1}(\lambda_1 P_i + \lambda_2 Q_i).$$

Then, we have the following result.

Result 1: Given (λ_1, λ_2), the process outputs of the IIIA controller are

$$Y_t^* - \tau = H_t + W_t, \qquad t \geq 3 \tag{23}$$

where W_t is given in (9), and

$$H_t = \sum_{i=0}^{t-2} h_{t-1,i}\eta_{t-1-i}. \tag{24}$$

The proof of Result 1 is given in Appendix III.

Result 1 claims that the deviation of the IIIA process output at tth run from the target τ can be written as the sum of H_t and W_t Note that since H_t in (23) is a random variable with zero mean, it is completely different from the nonrandom constant bias Γ_t in (7). Thus, we have $\mathrm{MSE}_t^* = \mathrm{Var}\,(Y_t^*)$. However, the derivation of MSE_t^* is tedious for general process disturbance because H_t and W_t are not statistically independent. Hence, in the following, we only focus on the derivation of MSE_t^* for the special case where $\eta_t \sim \mathrm{ARMA}(1,1)$.

Assume that $\eta_i \sim \mathrm{ARMA}(1,1)$ with parameters ϕ and θ for $i = 1, 2, \ldots$. That is

$$\eta_i = \phi\eta_{i-1} + \varepsilon_i - \theta\varepsilon_{i-1} \tag{25}$$

where $\{\varepsilon_i\}$ is a white noise series with mean zero and constant variance σ_ε^2. We further assume that $\eta_0 = 0$ and $\varepsilon_0 = 0$.

Via simple algebra, η_i in (25) can be written as

$$\eta_i = \varepsilon_i + (\phi - \theta)\sum_{j=0}^{i-2}\phi^j\varepsilon_{i-1-j}. \tag{26}$$

Plugging η_i in (24) into (26), we have

$$H_t = \sum_{i=1}^{t-1} Z_{t-1,t-i}\varepsilon_i$$

where

$$Z_{t-1,i} = h_{t-1,i-1} + (\phi - \theta)\sum_{j=0}^{i-2}\phi^j h_{t-1,i-2-j}$$
$$i = 1, \ldots, t-1. \tag{27}$$

Similarly, from (9), we have

$$W_t = \varepsilon_t + \sum_{i=1}^{t-1} G_i\varepsilon_{t-i}, \text{ for all } t \tag{28}$$

where

$$G_i = (E_i - E_{i-1}) + (\phi - \theta)\sum_{j=0}^{i-1}(E_j - E_{j-1})\phi^{i-1-j}$$
$$i = 1, 2, \ldots, t-1.$$

Note that $E_{-1} = 0$. Hence, it is straightforward to have the following result.

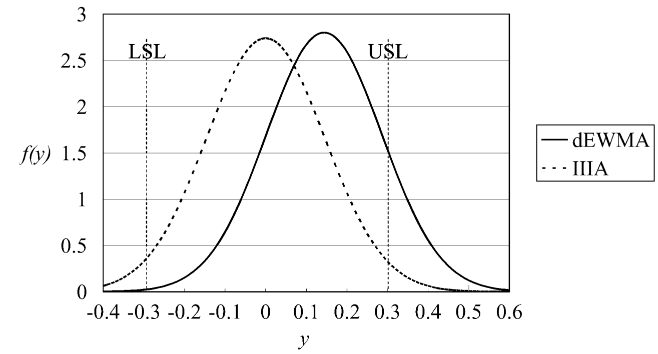


Fig. 3. Distribution of process outputs of the dEWMA and IIIA controllers when $t = 7$, $\tau = 0$, $\alpha = 4$, $\beta = 2$, $b = 2$, $\delta = 0.1$, $a_0 = 4.1$, $D_0 = 0$, $\lambda_1 = 0.1$, $\lambda_2 = 0.45$, $\eta_t \sim N(0, \sigma_\varepsilon^2 = 0.01)$, and the product's specification limit is $[\mathrm{LSL}, \mathrm{USL}] = [-0.3, 0.3]$.

Result 2: If $\eta_i \sim \mathrm{ARMA}(1,1)$, then

$$\begin{aligned}\mathrm{MSE}_t^* &= \mathrm{Var}(Y_t^*)\\ &= \sigma_\varepsilon^2 \left\{ 1 + \sum_{i=1}^{t-1} (Z_{t-1,i} + G_i)^2 \right\} \qquad \forall\, t \geq 3. \quad (29)\end{aligned}$$

The following example demonstrates the advantages of our proposed scheme.

Example 2—(Example 1 Revisited): Let the process target $\tau = 0$ and suppose that the process specification limits of the output [LSL, USL] are $[-0.3, 0.3]$. Now, we compare the process outputs of the IIIA controller with that of the dEWMA controller. Fig. 3 shows the distributions of the process output for the dEWMA and IIIA controllers when $t = 7$. From (8), (9), and (29), we have $\Gamma_7 = 0.1439$, $\mathrm{Var}(Y_7) = 0.0203$, and $\mathrm{Var}(Y_7^*) = 0.0212$. Note that the IIIA process mean meets the target (that is, $\Gamma_7^* = 0$). Let RR_t and RR_t^* denote the rework rate (RR) for the dEWMA and IIIA controllers at time t, respectively. Then, we have the following results:

$$\begin{cases} \mathrm{MSE}_7 = 0.0410, \mathrm{RR}_7 = 13.76\%, & \text{for the dEWMA scheme} \\ \mathrm{MSE}_7 = 0.0203, \mathrm{RR}_7 = 3.52\%, & \text{for the IIIA scheme} \end{cases}.$$

In other words, the IIIA controller reduces the rework rate (RR) by 74% $(= 13.76 - 3.52/13.76)$ and the MSE by 50% $(= 0.0410 - 0.0203/0.0410)$. ■

Now, we want to establish the necessary and sufficient conditions for guaranteeing that the MSE of IIIA is less than the MSE of dEWMA.

Result 3: Suppose that the following two conditions hold.

a) IIIA and dEWMA discount factors λ_1 and λ_2 are equal.

b) $\eta_i \sim \mathrm{ARMA}(1,1)$. Then

$$\mathrm{MSE}_t^* \leq \mathrm{MSE}_t \text{ if and only if } r_0(t) < \frac{\Gamma_t^2}{\sigma_\varepsilon^2} \text{ for all } t \geq 3 \quad (30)$$

where

$$r_0(t) = \sum_{i=1}^{t-1} \left(Z_{t-1,i}^2 + 2Z_{t-1,i} G_i \right).$$

The proof of Result 3 is given in Appendix IV.

We will call $\Gamma_t^2 / r_0(t)$ the IIIA *variance threshold* at time t. Hence, from (30), the MSE of IIIA is less than the MSE of dEWMA if σ_ε^2 is less than the value of the variance threshold.

Another way to investigate improvements in MSE is through the average MSE reduction ratio, which we define as

$$R = 1 - \frac{\sum_{t=3}^{N} \mathrm{MSE}_t^*}{\sum_{t=3}^{N} \mathrm{MSE}_t}.$$

Example 3 illustrates the effect of σ_ε on the average MSE reduction ratio.

Example 3: Suppose that the parameter settings in (1) and (2) are $\alpha = 4$, $\beta = 2$, and $\delta = 0.1$; the corresponding estimators of (α, β) are $(a_0, b) = (4.1, 2)$; and $\eta_t \sim \mathrm{ARMA}(1,1)$, where $-0.9 \leq \phi \leq 0.9$ and $-0.7 \leq \theta \leq 0.7$. Suppose in addition that $\sigma_\varepsilon = 0.025, 0.10$; $(\lambda_1, \lambda_2) = (0.1, 0.1), (0.25, 0.1), (0.25, 0.25), (0.5, 0.1), (0.5, 0.45)$; and $N = 30$. Table I shows that when $\sigma_\varepsilon = 0.025$, the average MSE reduction ratio R is positive. Specifically, the R values are larger than 0.491 for all combinations of (λ_1, λ_2) and (ϕ, θ) considered in our study. When $\sigma_\varepsilon = 0.10$, except for two cells in $(\lambda_1, \lambda_2) = (0.5, 0.1)$, all the R values are positive. Specifically, R is positive whenever $\phi \geq -0.50$. Consistent with Result 3, when ϕ, θ, λ_1, and λ_2 are fixed, we see that a smaller σ_ε leads to a larger value of R. ■

V. Simulation Study for Unknown ξ

When ξ is unknown, it is difficult to derive the analytical IIIA MSE results. Below, we use simulation to investigate the IIIA performance when ξ is unknown. Our comparison of the IIIA and the dEWMA controllers will include RLS. RLS [10], [11] is originally designed to update regression coefficients recursively. Hence, a control scheme that applies RLS (hereafter, we simply call it the RLS controller) could be used as a recursive tuning procedure to update the input recipe.

We now briefly state the square-root (SR) version of RLS controller [10] as follows: equation (1), the input and output model of a drifted process can be written as

$$Y_t = \mathbf{x}_t' \boldsymbol{\beta} + \eta_t$$

where

$$\begin{aligned}\mathbf{x}_t' &= [1 \quad u_{t-1} \quad t]\\ \boldsymbol{\beta}' &= [\alpha \quad \beta \quad \delta].\end{aligned}$$

Let $\hat{\boldsymbol{\beta}}_t (= [\hat{\alpha}_t \quad \hat{\beta}_t \quad \hat{\delta}_t])$ denote the LSE for $\boldsymbol{\beta}$ at time t. Then, we can write the iterative SR version of RLS algorithm (denoted by $\mathrm{RLS_{SR}}$) as follows. For $t = 1, 2, \ldots$, compute the following:

$$\begin{aligned}
&\text{Step 1)} \quad \mathbf{g}_t = \mathbf{S}_{t-1}' \mathbf{x}_t' \\
&\text{Step 2)} \quad f_t = \lambda(t) + \mathbf{g}_t' \mathbf{g}_t \\
&\text{Step 3)} \quad \alpha_t = \frac{1}{f_t + \sqrt{f_t \lambda(t)}} \\
&\text{Step 4)} \quad \mathbf{S}_t = \frac{1}{\sqrt{\lambda(t)}} \mathbf{S}_{t-1} \left(\mathbf{I} - \alpha_t \mathbf{g}_t \mathbf{g}_t' \right) \\
&\text{Step 5)} \quad \mathbf{M}_t = \mathbf{S}_{t-1} \mathbf{g}_t f_t^{-1} \\
&\text{Step 6)} \quad \hat{\boldsymbol{\beta}}_t = \hat{\boldsymbol{\beta}}_{t-1} + \mathbf{M}_t \left(Y_t - \mathbf{x}_t' \hat{\boldsymbol{\beta}}_{t-1} \right) \qquad (31)
\end{aligned}$$

where $\lambda(t)$ is called as the forgetting factor which can be suitably defined in [10] and the initial $\mathbf{S}_0$ is a given matrix (in our case below, $\mathbf{S}_0 = 10\mathbf{I}_{3\times 3}$).

TABLE I
COMPARISON OF THE AVERAGE IIIA AND dEWMA MSE REDUCTION RATIOS (R)
($N = 30, \alpha = 4, \beta = 1, \delta = 0.1, b = 1, \eta_t \sim \mathrm{ARMA}(1, 1)$, AND $\Gamma_0 = 3$)

		θ									
		-0.7		-0.3		0.0		0.3		0.7	
		σ_ε		σ_ε		σ_ε		σ_ε		σ_ε	
	(λ_1, λ_2)	0.025	0.10	0.025	0.10	0.025	0.10	0.025	0.10	0.025	0.10
ϕ -0.9	(0.10, 0.10)	0.997	0.956	0.994	0.904	0.989	0.840	0.982	0.762	0.971	0.647
	(0.25, 0.10)	0.989	0.846	0.975	0.691	0.957	0.540	0.933	0.396	0.892	0.239
	(0.25, 0.25)	0.975	0.694	0.943	0.460	0.903	0.289	0.851	0.163	0.772	0.053
	(0.50, 0.10)	0.930	0.381	0.840	0.125	0.741	0.007	0.631	-0.059	0.491	-0.106
	(0,50, 0.45)	0.960	0.578	0.901	0.321	0.831	0.186	0.750	0.104	0.639	0.042
-0.5	(0.10, 0.10)	0.998	0.964	0.997	0.960	0.997	0.946	0.995	0.924	0.992	0.883
	(0.25, 0.10)	0.992	0.878	0.990	0.859	0.987	0.813	0.981	0.745	0.970	0.633
	(0.25, 0.25)	0.981	0.756	0.978	0.717	0.970	0.635	0.956	0.528	0.932	0.378
	(0.50, 0.10)	0.949	0.492	0.937	0.414	0.913	0.293	0.876	0.167	0.812	0.033
	(0,50, 0.45)	0.972	0.676	0.964	0.610	0.949	0.506	0.925	0.391	0.884	0.259
0.0	(0.10, 0.10)	0.997	0.957	0.998	0.965	0.998	0.964	0.997	0.957	0.996	0.938
	(0.25, 0.10)	0.990	0.861	0.992	0.882	0.991	0.875	0.990	0.849	0.985	0.788
	(0.25, 0.25)	0.979	0.735	0.982	0.766	0.981	0.748	0.976	0.698	0.965	0.594
	(0.50, 0.10)	0.945	0.483	0.952	0.513	0.946	0.470	0.932	0.384	0.901	0.244
	(0,50, 0.45)	0.972	0.677	0.974	0.695	0.970	0.657	0.961	0.586	0.941	0.466
0.5	(0.10, 0.10)	0.996	0.939	0.997	0.958	0.998	0.965	0.998	0.966	0.997	0.959
	(0.25, 0.10)	0.987	0.822	0.991	0.869	0.992	0.886	0.992	0.886	0.990	0.856
	(0.25, 0.25)	0.973	0.688	0.981	0.757	0.983	0.780	0.983	0.771	0.977	0.713
	(0.50, 0.10)	0.936	0.449	0.952	0.530	0.956	0.547	0.953	0.516	0.937	0.412
	(0,50, 0.45)	0.970	0.667	0.977	0.723	0.978	0.727	0.974	0.695	0.964	0.610
0.9	(0.10, 0.10)	0.994	0.909	0.996	0.941	0.997	0.957	0.998	0.967	0.998	0.968
	(0.25, 0.10)	0.982	0.779	0.989	0.844	0.992	0.879	0.993	0.898	0.993	0.891
	(0.25, 0.25)	0.968	0.653	0.978	0.740	0.983	0.785	0.985	0.805	0.984	0.781
	(0.50, 0.10)	0.929	0.444	0.951	0.539	0.960	0.586	0.962	0.592	0.955	0.527
	(0,50, 0.45)	0.972	0.680	0.980	0.748	0.982	0.770	0.981	0.761	0.975	0.702

For the case of constant parameters, it was shown in [10] that $\lambda(t)$ should be chosen asymptotically to 1. In the following, two different settings of $\lambda(t)$ are considered.

1) If $\lambda(t) = 1$, we denote this kind of RLS scheme as $\mathrm{RLS_{SR}}(1)$.
2) If $\lambda(t) = \lambda_0\lambda(t-1)+(1-\lambda_0)$ with $\lambda_0 = 0.99$ and $\lambda(0) = 0.95$, we denote this kind of RLS scheme as $\mathrm{RLS_{SR}}(2)$.

Note that the above parameter settings in 2) has shown to work quite well for several low-order application ([10, p. 280]).

Next, we explain our simulation experiments. The parameters settings are $\alpha = 4$, $\beta = 0.9$, 1.0, 1.1 (corresponding to $\xi = 0.9$, 1.0, 1.1, respectively), $\delta = 0.1$, $\tau = 0$, $a_0 = 2$, $b = 1$, $D_0 = 0$, $\lambda_1 = 0.5$, and $\lambda_2 = 0.45$. Starting with the initial settings $\hat{\boldsymbol{\beta}}_0 = [\,a_0 \quad b \quad D_0\,]'$ and $\mathbf{S}_0 = 10\mathbf{I}_{3\times 3}$, the recipe is updated as follows:

$$u_t = \frac{\tau - \hat{\alpha}_t - \hat{\delta}_t(t+1)}{\hat{\beta}_t}, \qquad \forall\, t = 1, 2, \ldots \tag{32}$$

Note that $\xi(= \beta/b)$ is unknown and it needs to be estimated. By controlling the $100(1-\gamma)\%$ error bound to ensure that $(\xi - 1)$ is less than ε, we can determine the required sample size n such that ξ is probabilistically bounded by $(1-\varepsilon, 1+\varepsilon)$, which leads to the choice $\hat{\xi} = 1$. We use this value of ξ to estimate $\{A_t, B_t, C_t, O_t, P_t, Q_t\}$ for the IIIA control scheme. The simulation is based on two process disturbances, ARMA (1, 1) and ARIMA (1, 1, 1). 10,000 simulation trials with $N = 30$ are conducted for various combinations of ϕ, θ, and σ_ε.

Tables II–IV show the sample averages and (their sample standard errors) of the estimated MSE of the four controllers based on 10,000 simulation trials for $\xi = 0.9$, 1.0, 1.1, respectively. When σ_ε is large (especially when $\sigma_\varepsilon \geq 0.25$), the estimated MSE of process output of both $\mathrm{RLS_{SR}}(1)$ and $\mathrm{RLS_{SR}}(2)$ controllers may have a very large standard error (in which we simply denote it by the symbol X). The main reason is that with a larger σ_ε, both $\mathrm{RLS_{SR}}$ controllers will have a high probability of $\hat{\beta}_t \approx 0$. Note that from (32), at a specific time t, if $\hat{\beta}_t \approx 0$, then the corresponding process output will have a very large standard error. In addition, even under process disturbance follows a white series (that is, $(\phi, \theta) = (0, 0)$), the performance of both $\mathrm{RLS_{SR}}$ are still worse than the dEWMA and IIIA controllers.

The results shown in Tables II–IV establish the superior performance of the IIIA controller. For the case of nonstationary ARIMA (1, 1, 1) disturbance, the performance of the IIIA controller is significantly better than that of the dEWMA controller. Even for the case of stationary ARMA (1, 1) disturbance, the IIIA has a better performance except for few cases of a stationary ARMA (1, 1) disturbance with larger σ_ε (say, $\sigma_\varepsilon = 0.5$). The simulation results also matched with Result 3. In summary, we recommend the IIIA controller because of its overall robustness.

VI. CONCLUSION

For a SISO process with linear drift, this study proposes an IIIA controller that allows us to eliminate the nonrandom bias efficiently. We have stated conditions that are necessary and sufficient to ensure that the proposed IIIA controller outperforms the traditional dEWMA controller in terms of MSE.

TABLE II
SAMPLE AVERAGES (AND SAMPLE STANDARD ERRORS) OF THE ESTIMATED MSE OBTAINED FROM 10 000 SIMULATION TRIALS FOR $N = 30$, $\alpha = 4$, $\beta = 0.9$ ($\xi = 0.9$), $\delta = 0.1, \tau = 0, b = 1, a_0 = 2, D_0 = 0, \lambda_1 = 0.5$, AND $\lambda_2 = 0.45$

		η_t ~ARMA(1, 1)				η_t ~ARIMA(1, 1, 1)			
σ_ε	(ϕ,θ)	dEWMA	IIIA	$RLS_{SR}(1)$	$RLS_{SR}(2)$	dEWMA	IIIA	$RLS_{SR}(1)$	$RLS_{SR}(2)$
0.05	(0.0, 0.0)	0.22059 (0.00008)	**0.19575** **(0.00007)**	2.46206 (0.00568)	2.46424 (0.00568)	0.21896 (0.00010)	**0.19372** **(0.00009)**	2.46204 (0.00536)	2.46325 (0.00537)
	(0.5, 0.0)	0.21975 (0.00011)	**0.19467** **(0.00009)**	2.48385 (0.00647)	2.48591 (0.00648)	0.22049 (0.00013)	**0.19507** **(0.00011)**	2.49419 (0.00607)	2.49284 (0.00608)
	(0.0, -0.3)	0.22016 (0.00009)	**0.19518** **(0.00007)**	2.4674 (0.00586)	2.46954 (0.00586)	0.21965 (0.00012)	**0.19435** **(0.00010)**	2.47048 (0.00551)	2.47104 (0.00552)
	(0.5, 0.3)	0.22002 (0.00009)	**0.19507** **(0.00007)**	2.46488 (0.00576)	2.46701 (0.00577)	0.21929 (0.00011)	**0.19399** **(0.00009)**	2.46842 (0.00543)	2.46882 (0.00543)
	(0.5, -0.3)	0.22017 (0.00013)	**0.19502** **(0.00011)**	2.52127 (0.00777)	2.52325 (0.00778)	0.22253 (0.00016)	**0.19699** **(0.00014)**	2.54010 (0.00729)	2.53637 (0.00730)
0.1	(0.0, 0.0)	0.23461 (0.00016)	**0.21139** **(0.00014)**	2.70434 (0.01381)	2.70693 (0.01383)	0.22817 (0.00021)	**0.20335** **(0.00018)**	2.70261 (0.01310)	2.70129 (0.01311)
	(0.5, 0.0)	0.23129 (0.00022)	**0.20708** **(0.00018)**	2.81502 (0.01711)	2.81716 (0.01713)	0.23433 (0.00027)	**0.20879** **(0.00023)**	2.85330 (0.01621)	2.84157 (0.01623)
	(0.0, -0.3)	0.23294 (0.00018)	**0.20915** **(0.00016)**	2.73038 (0.01450)	2.73280 (0.01452)	0.23097 (0.00023)	**0.20590** **(0.00020)**	2.74077 (0.01373)	2.73681 (0.01374)
	(0.5, 0.3)	0.23235 (0.00018)	**0.20867** **(0.00015)**	2.71840 (0.01421)	2.72079 (0.01423)	0.22951 (0.00022)	**0.20442** **(0.00019)**	2.73067 (0.01346)	2.72604 (0.01348)
	(0.5, -0.3)	0.23303 (0.00027)	**0.20850** **(0.00022)**	3.02535 (0.02396)	3.02735 (0.02399)	0.24249 (0.00033)	**0.21647** **(0.00029)**	3.09464 (0.02283)	3.07322 (0.02286)
0.25	(0.0, 0.0)	0.33319 (0.00051)	**0.32122** **(0.00051)**	X	X	0.29318 (0.00056)	**0.27121** **(0.00049)**	X	X
	(0.5, 0.0)	0.31259 (0.00060)	**0.29442** **(0.00053)**	X	X	0.33175 (0.00078)	**0.30531** **(0.00070)**	X	X
	(0.0, -0.3)	0.32283 (0.00054)	**0.30727** **(0.00051)**	X	X	0.31075 (0.00066)	**0.28721** **(0.00058)**	X	X
	(0.5, 0.3)	0.31912 (0.00052)	**0.30424** **(0.00049)**	X	X	0.30162 (0.00062)	**0.27793** **(0.00054)**	X	X
	(0.5, -0.3)	0.32355 (0.00074)	**0.30338** **(0.00064)**	X	X	0.38284 (0.00104)	**0.35342** **(0.00094)**	X	X
0.5	(0.0, 0.0)	**0.68582** **(0.00156)**	0.71394 (0.00172)	X	X	0.52614 (0.00138)	**0.51427** **(0.00129)**	X	X
	(0.5, 0.0)	**0.60363** **(0.00154)**	0.60691 (0.00150)	X	X	0.68052 (0.00218)	**0.65080** **(0.00203)**	X	X
	(0.0, -0.3)	**0.64450** **(0.00153)**	0.65825 (0.00158)	X	X	0.59654 (0.00172)	**0.57837** **(0.00161)**	X	X
	(0.5, 0.3)	**0.62962** **(0.00145)**	0.64608 (0.00152)	X	X	0.55995 (0.00157)	**0.54122** **(0.00146)**	X	X
	(0.5, -0.3)	0.64761 (0.00191)	**0.64289** **(0.00179)**	X	X	0.88493 (0.00315)	**0.84336** **(0.00296)**	X	X

The proposed controller is studied under the assumption that the process follows a SISO model. In practice, however, manufacturing processes follow multiple-input–multiple-output (MIMO) models (refer to [13]). Thus, further investigation is needed to determine to what extent that the proposed procedure can be applied to MIMO processes.

APPENDIX I
PROOF OF LEMMA 1

Substitute (1) and (3) into (4) and (5) to obtain the following equations:

$$\begin{aligned} a_t &= \lambda_1[\alpha + \tau(\xi - 1)] + (1 - \lambda_1\xi)a_{t-1} \\ &\quad - \lambda_1(\xi - 1)D_{t-1} + \lambda_1\delta t + \lambda_1\eta_t \\ D_t &= \lambda_2[\alpha + \tau(\xi - 1)] - \lambda_2\xi a_{t-1} \\ &\quad + (1 - \lambda_2\xi)D_{t-1} + \lambda_2\delta t + \lambda_2\eta_t. \end{aligned}$$

Recursive substitution of the above two equations yields the following:

$$\begin{aligned} a_t &= A_t(\alpha + \tau(\xi - 1)) + B_t a_0 + C_t D_0 \\ &\quad + \delta\sum_{i=0}^{t}(\lambda_1 B_i + \lambda_2 C_i)(t - i) \\ &\quad + \sum_{i=0}^{t}(\lambda_1 B_i + \lambda_2 C_i)\eta_{t-i}, \\ D_t &= O_t(\alpha + \tau(\xi - 1)) + P_t a_0 + Q_t D_0 \\ &\quad + \delta\sum_{i=0}^{t}(\lambda_1 P_i + \lambda_2 Q_i)(t - i) \\ &\quad + \sum_{i=0}^{t}(\lambda_1 P_i + \lambda_2 Q_i)\eta_{t-i}. \end{aligned}$$

These are equivalent to (14) and (15). ■

APPENDIX II
PROOF OF LEMMA 2

From the definitions of A_t, O_t, Λ_t, K_t, Φ_t, and L_t in Section III-A, no term on the right-hand side (RHS) of (14) asnd (15), except for K_t and L_t, contain a_0. Hence, changing a_0 to a_0^*, we can easily express the updating filter $a_t^*(D_t^*)$ as a function of $a_t(D_t)$ in (18) and (19) via the following two equations:

$$\begin{aligned} K_t &= A_t\tau(\xi - 1) + B_t a_0 + C_t D_0 \\ L_t &= O_t\tau(\xi - 1) + P_t a_0 + Q_t D_0. \end{aligned}$$ ■

TABLE III
SAMPLE AVERAGES (AND SAMPLE STANDARD ERRORS) OF THE ESTIMATED MSE OBTAINED FROM 10000 SIMULATION TRIALS FOR $N = 30$, $\alpha = 4$, $\beta = 1.0$ ($\xi = 1.0$), $\delta = 0.1$, $\tau = 0$, $b = 1$, $a_0 = 2$, $D_0 = 0$, $\lambda_1 = 0.5$, AND $\lambda_2 = 0.45$

		η_t~ARMA(1, 1)				η_t~ARIMA(1, 1, 1)			
σ_ε	(ϕ,θ)	dEWMA	IIIA	$RLS_{SR}(1)$	$RLS_{SR}(2)$	dEWMA	IIIA	$RLS_{SR}(1)$	$RLS_{SR}(2)$
0.05	(0.0, 0.0)	0.17795 (0.00007)	**0.15913** **(0.00007)**	1.45366 (0.00290)	1.45472 (0.00291)	0.17589 (0.00009)	**0.15659** **(0.00007)**	1.45465 (0.00266)	1.45473 (0.00267)
	(0.5, 0.0)	0.17684 (0.00010)	**0.15770** **(0.00008)**	1.46243 (0.00324)	1.46334 (0.00325)	0.17703 (0.00011)	**0.15753** **(0.00009)**	1.47414 (0.00294)	1.47161 (0.00294)
	(0.0, -0.3)	0.17736 (0.00008)	**0.15834** **(0.00007)**	1.45574 (0.00297)	1.45675 (0.00298)	0.17643 (0.00010)	**0.15704** **(0.00008)**	1.45993 (0.00271)	1.45934 (0.00271)
	(0.5, 0.3)	0.17726 (0.00008)	**0.15830** **(0.00007)**	1.45462 (0.00293)	1.45562 (0.00293)	0.17610 (0.00009)	**0.15672** **(0.00008)**	1.45922 (0.00267)	1.45847 (0.00268)
	(0.5, -0.3)	0.17712 (0.00012)	**0.15788** **(0.00010)**	1.47810 (0.00384)	1.47889 (0.00384)	0.17870 (0.00013)	**0.15908** **(0.00011)**	1.49884 (0.00348)	1.49385 (0.00347)
0.1	(0.0, 0.0)	0.19298 (0.00016)	**0.17615** **(0.00015)**	1.55651 (0.00653)	1.55772 (0.00654)	0.18483 (0.00018)	**0.16601** **(0.00015)**	1.55936 (0.00601)	1.55665 (0.00602)
	(0.5, 0.0)	0.18858 (0.00020)	**0.17046** **(0.00017)**	1.59650 (0.00764)	1.59713 (0.00764)	0.18937 (0.00022)	**0.16979** **(0.00019)**	1.64066 (0.00700)	1.62721 (0.00699)
	(0.0, -0.3)	0.19065 (0.00017)	**0.17301** **(0.00015)**	1.56574 (0.00676)	1.56675 (0.00677)	0.18697 (0.00020)	**0.16782** **(0.00016)**	1.58112 (0.00619)	1.57569 (0.00620)
	(0.5, 0.3)	0.19028 (0.00017)	**0.17285** **(0.00015)**	1.56081 (0.00664)	1.56181 (0.00665)	0.18565 (0.00019)	**0.16653** **(0.00016)**	1.57784 (0.00609)	1.57175 (0.00610)
	(0.5, -0.3)	0.18974 (0.00024)	**0.17117** **(0.00020)**	1.67156 (0.00973)	1.67177 (0.00974)	0.19608 (0.00027)	**0.17601** **(0.00023)**	1.74890 (0.00895)	1.72526 (0.00894)
0.25	(0.0, 0.0)	0.29866 (0.00052)	**0.29574** **(0.00056)**	3.41918 (0.23252)	3.42332 (0.23154)	0.24785 (0.00049)	**0.23244** **(0.00043)**	3.40800 (0.23208)	3.38687 (0.23108)
	(0.5, 0.0)	0.27128 (0.00056)	**0.26021** **(0.00051)**	X	X	0.27625 (0.00065)	**0.25608** **(0.00057)**	X	X
	(0.0, -0.3)	0.28413 (0.00052)	**0.27606** **(0.00051)**	X	X	0.26126 (0.00056)	**0.24378** **(0.00049)**	X	X
	(0.5, 0.3)	0.28178 (0.00050)	**0.27509** **(0.00051)**	4.12139 (0.69527)	4.11221 (0.68205)	0.25298 (0.00052)	**0.23571** **(0.00046)**	4.18965 (0.69753)	4.13485 (0.68427)
	(0.5, -0.3)	0.27856 (0.00067)	**0.26463** **(0.00059)**	X	X	0.31821 (0.00086)	**0.29499** **(0.00076)**	X	X
0.5	(0.0, 0.0)	**0.67664** **(0.00165)**	0.72341 (0.00192)	X	X	0.47361 (0.00123)	**0.47030** **(0.00118)**	X	X
	(0.5, 0.0)	**0.56726** **(0.00148)**	0.58132 (0.00149)	X	X	0.58726 (0.00183)	**0.56493** **(0.00169)**	X	X
	(0.0, -0.3)	**0.61860** **(0.00153)**	0.64474 (0.00164)	X	X	0.52734 (0.00150)	**0.51571** **(0.00140)**	X	X
	(0.5, 0.3)	**0.60917** **(0.00148)**	0.64083 (0.00164)	X	X	0.49415 (0.00136)	**0.48343** **(0.00127)**	X	X
	(0.5, -0.3)	**0.59651** **(0.00177)**	0.59905 (0.00168)	X	X	0.75513 (0.00261)	**0.72058** **(0.00243)**	X	X

APPENDIX III
PROOF OF RESULT 1

Let X be a random variable and define $(X)_R = X - \mathrm{E}(X)$. From (14)–(16) and (20), we have $(\hat{\alpha}_t) = \alpha$, and

$$(\hat{\alpha}_t)_R = \frac{(\Phi_t\zeta_t - \Lambda_t\upsilon_t)}{(A_t\Phi_t - O_t\Lambda_t)}. \tag{A1}$$

Similarly, (14), (15), (17), and (21) yield $\mathrm{E}(\hat{\delta}_t) = \delta$, and

$$(\hat{\delta}_t)_R = \frac{(O_t\zeta_t - A_t\upsilon_t)}{(O_t\Lambda_t - A_t\Phi_t)}. \tag{A2}$$

From (12), (16), (17), and (20)–(22), we have

$$\hat{a}_0^*(t) = \frac{a_0^*(t) + (E(\hat{\alpha}_t)_R + (E_t + F_t)(\hat{\delta}_t)_R)}{(\xi E_t)}.$$

Thus, we obtain

$$(\hat{a}_0^*(t))_R = \frac{(E_t(\hat{\alpha}_t)_R + (E_t + F_t)(\hat{\delta})_R)}{(\xi E_t)}. \tag{A3}$$

We use equations (A1)–(A3) as follows.

Let $\Gamma_t\left(\hat{a}_0^*(t-1)\right)$ denote the nonrandom part of the dEWMA process output at the tth run with the initial intercept $\hat{a}_0^*(t-1)$. Then, from (7), (8), and (10), we have

$$\begin{aligned}(Y_t^* - \tau) &= \Gamma_t\left(\hat{a}_0^*(t-1)\right) + W_t \\ &= -(\hat{\alpha}_{t-1})_R E_{t-1} - (\hat{\delta}_{t-1})_R(F_{t-1} + E_{t-1}) + W_t.\end{aligned}$$

From (A2) and (A3), we have

$$\begin{aligned}H_t = &\frac{\Phi_{t-1}\zeta_{t-1} - \Lambda_{t-1}\upsilon_{t-1}}{O_{t-1}\Lambda_{t-1} - A_{t-1}\Phi_{t-1}}E_{t-1} \\ &- \frac{O_{t-}\zeta_{t-1} - A_{t-1}\upsilon_{t-1}}{O_{t-1}\Lambda_{t-1} - A_{t-1}\Phi_{t-1}}(F_{t-1} + E_{t-1}).\end{aligned}$$

Substituting the expressions for ζ_{t-1} and υ_{t-1} in 2) and 4) of Section III-A into the above equation, we obtain Result 2. ■

TABLE IV
SAMPLE AVERAGES (AND SAMPLE STANDARD ERRORS) OF THE ESTIMATED MSE OBTAINED FROM 10 000 SIMULATION TRIALS FOR $N = 30$, $\alpha = 4$, $\beta = 1.1$ ($\xi = 1.1$), $\delta = 0.1$, $\tau = 0$, $b = 1$, $a_0 = 2$, $D_0 = 0$, $\lambda_1 = 0.5$, AND $\lambda_2 = 0.45$

		η_t ~ARMA(1, 1)				η_t ~ARIMA(1, 1, 1)			
σ_ε	(ϕ, θ)	dEWMA	IIIA	$RLS_{SR}(1)$	$RLS_{SR}(2)$	dEWMA	IIIA	$RLS_{SR}(1)$	$RLS_{SR}(2)$
0.05	(0.0, 0.0)	0.14339 (0.00007)	**0.13207 (0.00008)**	0.92781 (0.00172)	0.92840 (0.00172)	0.14091 (0.00008)	**0.12897 (0.00007)**	0.92930 (0.00153)	0.92891 (0.00153)
	(0.5, 0.0)	0.14200 (0.00009)	**0.13026 (0.00008)**	0.93214 (0.00189)	0.93258 (0.00189)	0.14171 (0.00009)	**0.12957 (0.00007)**	0.94450 (0.00165)	0.94147 (0.00165)
	(0.0, -0.3)	0.14261 (0.00008)	**0.13101 (0.00008)**	0.92879 (0.00175)	0.92933 (0.00175)	0.14129 (0.00008)	**0.12925 (0.00007)**	0.93352 (0.00154)	0.93247 (0.00154)
	(0.5, 0.3)	0.14259 (0.00007)	**0.13108 (0.00008)**	0.92814 (0.00172)	0.92868 (0.00172)	0.14102 (0.00008)	**0.12898 (0.00007)**	0.93328 (0.00152)	0.93206 (0.00152)
	(0.5, -0.3)	0.14213 (0.00011)	**0.13025 (0.00009)**	0.94040 (0.00221)	0.94072 (0.00221)	0.14309 (0.00011)	**0.13082 (0.00009)**	0.96199 (0.00193)	0.95647 (0.00193)
0.1	(0.0, 0.0)	0.15960 (0.00015)	**0.15079 (0.00017)**	0.98449 (0.00373)	0.98519 (0.00373)	0.14970 (0.00016)	**0.13838 (0.00014)**	0.98958 (0.00333)	0.98631 (0.00333)
	(0.5, 0.0)	0.15404 (0.00018)	**0.14355 (0.00017)**	1.00320 (0.00422)	1.00328 (0.00423)	0.15289 (0.00019)	**0.14075 (0.00016)**	1.04963 (0.00375)	1.03541 (0.00373)
	(0.0, -0.3)	0.15649 (0.00016)	**0.14654 (0.00016)**	0.98862 (0.00382)	0.98911 (0.00382)	0.15123 (0.00017)	**0.13949 (0.00014)**	1.00633 (0.00339)	1.00031 (0.00339)
	(0.5, 0.3)	0.15640 (0.00016)	**0.14682 (0.00016)**	0.98590 (0.00376)	0.98637 (0.00376)	0.15007 (0.00016)	**0.13842 (0.00014)**	1.00513 (0.00334)	0.99844 (0.00334)
	(0.5, -0.3)	0.15459 (0.00022)	**0.14349 (0.00019)**	1.04000 (0.00519)	1.03960 (0.00519)	0.15842 (0.00023)	**0.14577 (0.00019)**	1.11959 (0.00463)	1.09497 (0.00460)
0.25	(0.0, 0.0)	0.27345 (0.00053)	**0.28233 (0.00064)**	1.57998 (0.02783)	1.58076 (0.02773)	0.21165 (0.00043)	**0.20465 (0.00041)**	1.58557 (0.02582)	1.56207 (0.02587)
	(0.5, 0.0)	0.23878 (0.00052)	**0.23698 (0.00052)**	X	X	0.23161 (0.00055)	**0.21942 (0.00048)**	X	X
	(0.0, -0.3)	0.25409 (0.00051)	**0.25574 (0.00055)**	1.63829 (0.03037)	1.63879 (0.03043)	0.22128 (0.00049)	**0.21159 (0.00043)**	1.71778 (0.02868)	1.67695 (0.02873)
	(0.5, 0.3)	0.25351 (0.00050)	**0.25749 (0.00056)**	1.64769 (0.04719)	1.63339 (0.03927)	0.21399 (0.00045)	**0.20492 (0.00041)**	1.70275 (0.03432)	1.65781 (0.03441)
	(0.5, -0.3)	0.24228 (0.00061)	**0.23659 (0.00057)**	5.26600 (1.11889)	5.27210 (1.10700)	0.26615 (0.00071)	**0.25081 (0.00063)**	5.58490 (1.11943)	6.77461 (1.72316)
0.5	(0.0, 0.0)	**0.68062 (0.00177)**	0.75281 (0.00221)	X	X	0.43354 (0.00113)	**0.44194 (0.00113)**	X	X
	(0.5, 0.0)	**0.54204 (0.00144)**	0.57132 (0.00155)	X	X	0.51336 (0.00155)	**0.50096 (0.00144)**	X	X
	(0.0, -0.3)	**0.60325 (0.00155)**	0.64638 (0.00177)	X	X	0.47209 (0.00132)	**0.46967 (0.00125)**	X	X
	(0.5, 0.3)	**0.60089 (0.00153)**	0.65340 (0.00184)	X	X	0.44290 (0.00120)	**0.44301 (0.00115)**	X	X
	(0.5, -0.3)	0.55609 (0.00165)	**0.56969 (0.00163)**	X	X	0.65154 (0.00219)	**0.62649 (0.00203)**	X	X

APPENDIX IV
PROOF OF RESULT 3

From (7), (8), (28), and (29), then we have

$$\mathrm{MSE}(Y_t^*;\tau) < \mathrm{MSE}(Y_t;\tau) \text{ if and only if } \sum_{i=1}^{t-1}\left(Z_{t-1,i}^2 + 2Z_{t-1,i}G_i\right) < \frac{\Gamma_t^2}{\sigma_\varepsilon^2}.$$

This gives Result 3. ■

REFERENCES

[1] G. E. P. Box, G. M. Jenkins, and G. C. Reinsel, *Time Series Analysis: Forecasting and Control*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall, 1994.

[2] G. E. P. Box and A. Luceño, *Statistical Control by Monitoring and Feedback Adjustment*. New York: Wiley, 1997.

[3] S. W. Butler and J. A. Stefani, "Supervisory run-to-run control of a polysilicon gate etch using *in situ* ellipsometry," *IEEE Trans. Semicond. Manuf.*, vol. 7, pp. 193–201, 1994.

[4] A. Chen and R. S. Guo, "Age-based double EWMA controller and its application to CMP processes," *IEEE Trans. Semicond. Manuf.*, vol. 14, no. 1, pp. 11–19, Feb. 2001.

[5] E. Del Castillo and A. Hurwitz, "Run to run process control: A literature review and some extensions," *J. Qual. Technol.*, vol. 29, no. 2, pp. 184–196, 1997.

[6] E. Del Castillo, "Long run and transient analysis of a double EWMA feedback controller," *IIE Trans.*, vol. 31, no. 12, pp. 1157–1169, 1999.

[7] ——, *Statistical Process Adjustment for Quality Control*. New York: Wiley, 2002.

[8] M. Hankinson, T. Vincent, K. B. Irani, and P. P. Khargonekar, "Integrated real-time and run-to-run control of etch depth in reactive ion etching," *IEEE Trans. Semicond. Manuf.*, vol. 10, no. 1, pp. 121–130, Feb. 1997.

[9] A. Ingolfsson and E. Sachs, "Stability and sensitivity of an EWMA controller," *J. Qual. Technol.*, vol. 25, no. 4, pp. 271–287, 1993.

[10] L. Ljung and T. Soderstrom, *Theory and Practice of Recursive Identification*. Cambridge: MIT Press, 1983, ch. A.

[11] D. S. G. Pollock, *A Handbook of Time-Series Analysis, Signal Processing and Dynamics*. New York: Academic, 1999.

[12] S. T. Tseng, R. J. Chou, and S. P. Lee, "Statistical design of a double EWMA controller," *Appl. Stochastic Models Bus. Industry*, vol. 18, no. 3, pp. 313–322, 2002.

[13] ——, "A study of multivariate EWMA controller," *IIE Trans.*, vol. 34, no. 6, pp. 541–549, 2002.

[14] F. Tsung, H. Wu, and V. N. Nair, "On the efficiency and robustness of discrete proportional-integral control schemes," *Technometrics*, vol. 40, no. 3, pp. 214–222, 1998.

[15] F. Tsung and J. Shi, "Integrated design of run-to-run PID controller and SPC monitoring for process disturbance rejection," *IIE Trans.*, vol. 31, no. 6, pp. 517–527, 1999.

[16] F. Tsung, J. Shi, and C. F. J. Wu, "Joint monitoring of PID-controlled processes," *J. Qual. Technol.*, vol. 31, no. 3, pp. 275–285, 1999.

Sheng-Tsaing Tseng (M'92) received the B.S. degree in business mathematics from Soochow University, Taipei, Taiwan, the M.S. degree in applied mathematics from Tsing-Hua University, Hsin-Chu, Taiwan, and the Ph.D. degree in management science from Tamkang University, Taipei, Taiwan.

Currently, he is a Professor in the Institute of Statistics, National Tsing-Hua University, Hsin-Chu, Taiwan. His articles have appeared in numerous technical journals. His research interests include quality and productivity improvement, reliability lifetime analysis, and statistical decision methodology.

Prof. Tseng received the Outstanding Research Award from the National Science Council (NSC), Taiwan, R.O.C., in 1993, 1999, and 2004, respectively. He is an elected member of ISI, and a member of ASQ.

Wheyming (Tina) Song received the B.S. degree in statistics and the M. S. degree in industrial management from Cheng-Kung University, Tainan, Taiwan, in 1979, the M.S. degrees in applied mathematics and industrial engineering from the University of Pittsburgh, Pittsburgh, PA, both in 1984, and the Ph.D. degree from the School of Industrial Engineering, Purdue University, West Lafayette, IN, in 1989.

She joined the National Tsing-Hua University, Hsin-Chu, Taiwan, in 1990, after spending one year as a Visiting Assistant Professor at the Department of Industrial Engineering, Purdue University. Currently, she is a Professor in the Department of Industrial Engineering, National Tsing-Hua University. Her research interests are applied operations research, probability and statistics, and stochastic simulation, including input modeling, output analysis, variance reduction, and ranking and selection

Prof. Song received the Outstanding Teaching Award from the National Tsing-Hua University in 1993. Her research honors include the 1988 Omega Rho Student Paper Award, the 1990 IIE Doctoral Dissertation Award, and the 1996 Outstanding Research Award from the National, Science Council (NSC), R.O.C.

Yu-Chi Chang received the B.S. degree in mathematics and the M.S. degree in industrial engineering from the National Tsing-Hua University, Hsin-Chu, Taiwan, in 1987 and 1989, respectively. Currently, he is working towards the Ph.D. degree in industrial engineering at the National Tsing-Hua University.

After graduation, from 1989 to 1993, he served as an Associate Researcher in the Research and Service Organization (ERSO), Industrial Technology Research Institute (ITRI), Hsin-Chu, Taiwan. From 1993 to 1999, he was with Nan-Kai Institute of Technology. In 1999, he was with the Department of Industrial Engineering and Management, Chienkuo Technology University, Changhua, Taiwan. His research interests include run-to-run (R2R) process control, probability and statistics, and stochastic simulation